

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:20:03 ; Search time 23.55 Seconds
(without alignments)
999.993 Million cell updates/sec

Title: US-09-818-066-34
Perfect score: 871
Sequence: 1 MGQHPAKSMVRRIEGGEIL.....PLYQSEPAVPVTKTPLKKK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 871 | 100.0 | 329 | 12 | O92935 |
| 2 | 866 | 99.4 | 364 | 12 | O66399 |
| 3 | 785.5 | 90.2 | 330 | 12 | O72885 |
| 4 | 779.5 | 89.5 | 330 | 12 | O66405 |
| 5 | 779.5 | 89.5 | 366 | 12 | O66404 |
| 6 | 724 | 83.1 | 329 | 12 | O9WFB3 |
| 7 | 724 | 83.1 | 329 | 12 | O9WFA9 |
| 8 | 724 | 83.1 | 329 | 12 | O9WFA3 |
| 9 | 717 | 82.3 | 329 | 12 | O9WFB6 |
| 10 | 714 | 82.0 | 329 | 12 | O9WFA6 |
| 11 | 602.5 | 69.2 | 327 | 12 | O67852 |
| 12 | 98.5 | 11.3 | 418 | 2 | O9A4Y7 |
| 13 | 98 | 11.3 | 1172 | 4 | O9Y4F2 |
| 14 | 96 | 11.0 | 548 | 4 | O9BTI7 |
| 15 | 96 | 11.0 | 559 | 4 | O9BYJ9 |
| 16 | 95.5 | 11.0 | 315 | 11 | O55150 |
| 17 | 95.5 | 11.0 | 633 | 10 | O65655 |
| 18 | 95.5 | 11.0 | 1339 | 11 | O35788 |
| 19 | 94.5 | 10.8 | 542 | 12 | O84357 |

| | | | | | |
|----|------|------|------|----|--------|
| 20 | 94 | 10.8 | 1260 | 4 | O9UGY9 |
| 21 | 93.5 | 10.7 | 907 | 4 | P78344 |
| 22 | 93.5 | 10.7 | 907 | 6 | P79398 |
| 23 | 92.5 | 10.6 | 269 | 10 | O9FUR7 |
| 24 | 92.5 | 10.6 | 989 | 11 | O9JLE9 |
| 25 | 92 | 10.6 | 539 | 10 | O9M2B7 |
| 26 | 92 | 10.6 | 980 | 12 | O92809 |
| 27 | 92 | 10.6 | 981 | 12 | O39477 |
| 28 | 91.5 | 10.5 | 299 | 10 | O49201 |
| 29 | 91.5 | 10.5 | 1386 | 4 | O9COA3 |
| 30 | 91 | 10.4 | 197 | 4 | O9NX79 |
| 31 | 90.5 | 10.4 | 926 | 3 | O13305 |
| 32 | 90 | 10.3 | 407 | 2 | O53879 |
| 33 | 89.5 | 10.3 | 10.3 | 11 | O62448 |
| 34 | 89 | 10.2 | 283 | 5 | O20349 |
| 35 | 89 | 10.2 | 593 | 11 | O55153 |
| 36 | 89 | 10.2 | 593 | 11 | O99MG2 |
| 37 | 89 | 10.2 | 616 | 11 | O55151 |
| 38 | 89 | 10.2 | 654 | 11 | O55152 |
| 39 | 89 | 10.2 | 654 | 11 | O99MG1 |
| 40 | 88.5 | 10.2 | 493 | 10 | O9LT85 |
| 41 | 88.5 | 10.2 | 1315 | 10 | O9SPM0 |
| 42 | 88 | 10.1 | 337 | 11 | P97268 |
| 43 | 87.5 | 10.0 | 396 | 4 | O9H5Q4 |
| 44 | 87.5 | 10.0 | 413 | 5 | P90973 |
| 45 | 87.5 | 10.0 | 453 | 5 | O02592 |

ALIGNMENTS

RESULT 1

O92935
ID O92935 PRELIMINARY; PRT; 329 AA.
AC O92935;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PRES ANTIGEN
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALBERTA;
RA Fischer K.P., Stickney J., Tipples G.A., Tyrrell D.L.J.;
RT "Cloning, sequencing and sequence comparison of a Canadian isolate of duck hepatitis B virus."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047045; AAC06355.1;
DR InterPro: IPR000349; Hepadnavir_surfa9.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 329 AA; 36361 MW; 46E4ACAF4995147 CRC64;

Query Match 100.0%; Score 871; DB 12; Length 329;
Best Local Similarity 100.0%; Pred. No. 6.4e-73;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTITWSGKFFPTLDHVDHVTMEINTLQ 60
Db 2 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTITWSGKFFPTLDHVDHVTMEINTLQ 61
Qy 61 NOGAWPAGAGRVGLSNPTPOEIPQPTWPEEDQKARAFRYQERPEPTTTPPSPP 120
Db 62 NOGAWPAGAGRVGLSNPTPOEIPQPTWPEEDQKARAFRYQERPEPTTTPPSPP 121
Qy 121 QWKLPQGGDDPLLGNOSLLETHPLYSQSEPAVPVTKTPLKKK 161
Db 122 QWKLPQGGDDPLLGNOSLLETHPLYSQSEPAVPVTKTPLKKK 162

RESULT 2

```
066399
ID Q66399 PRELIMINARY; PRT; 364 AA.
AC Q66399;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE SURFACE ANTIGEN.
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIAN;
RA Munshi A., Panda S.K.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74623; CAA52699.1; -.
DR InterPro: IPR000349; Hepadnavir_surfa9.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 364 AA; 40385 MW; E2E27FB4E4775C19 CRC64;

Query Match 99.4%; Score 866; DB 12; Length 364;
Best Local Similarity 99.4%; Pred. No. 2.1e-72;
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|||||
DB 37 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 96
|||||

QY 61 NOGAWPAGARRVGLSNPTPOETPQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 97 NOGAWPAGARRVGLSNPTPOETPQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 156
|||||

QY 121 QWKLPQDDPDLGNQSLLETHPLHYQSEPAVPVVIKTPPLKKK 161
|||||
DB 157 QWKLPQDDPDLGNQSLLETHPLHYQSEPAVPVVIKTPPLKKK 197
|||||

RESULT 3
ID 072885 PRELIMINARY; PRT; 330 AA.
AC 072885;
DT 01-AUG-1998 (TREMELrel. 07, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE SURFACE PROTEIN.
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AUSTRALIAN DHBV;
RA Triyatni M., Qiao M., Ey P., Burrell C., Jilbert A.R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ006350; CAA06988.1; -.
DR InterPro: IPR000349; Hepadnavir_surfa9.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 330 AA; 36789 MW; EDED4F42373ADA99 CRC64;

Query Match 90.2%; Score 785.5; DB 12; Length 330;
Best Local Similarity 88.9%; Pred. No. 5.4e-65;
Matches 144; Conservative 10; Mismatches 7; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|||||
DB 1 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|||||

QY 61 NOGAWPAGARRVGLSNPTPOETPQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 61 QOGAWPAGARRGLTNPAQEPQPPQWTPPEEDQKAREAFRRYQERPPETTTIPTST 120
|||||

QY 121 QWKLPQDDPDLGNQSLLETHPLHYQSEPAVPVVIKTPPLKKK 161
|||||
```

```
DB 121 PKWLPQDDPDLLENKSLLETHPLHYQNPPEAVPVVIKTPPLKKK 162
|||||

RESULT 4
ID Q66405 PRELIMINARY; PRT; 330 AA.
AC Q66405;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE PRE-S. PROTEIN.
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DHBVQCA34;
RA Schneider R.;
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DHBVQCA34;
RA Tong S., Mattes F., Blum H.E., Fernholz D., Schneider R., Will H.;
RL Nucleic Acids Res. 0:0-0(0).
DR EMBL; X60213; CAA42771.1; -.
DR InterPro: IPR000349; Hepadnavir_surfa9.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 330 AA; 36959 MW; 7CE142013BB8D9F4 CRC64;

Query Match 89.5%; Score 779.5; DB 12; Length 330;
Best Local Similarity 88.3%; Pred. No. 1.9e-64;
Matches 143; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|||||
DB 1 MGQHPAKSMVDVRIEGEILLNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|||||

QY 61 NOGAWPAGARRVGLSNPTPOETPQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 61 QOGAWPAGARRGLTNPTPHETPQWTPPEEDQKAREAFRRYQERPPETTTIPTST 120
|||||

QY 121 QWKLPQDDPDLGNQSLLETHPLHYQSEPAVPVVIKTPPLKKK 161
|||||
DB 121 PKWLPQDDPDLLENKSLLETHPLHYQNPPEAVPVVIKTPPLKKK 162
|||||

RESULT 5
ID Q66404 PRELIMINARY; PRT; 366 AA.
AC Q66404;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE PRE-S/S.
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DHBVQCA34;
RA Schneider R.;
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DHBVQCA34;
RA Tong S., Mattes F., Blum H.E., Fernholz D., Schneider R., Will H.;
RL Nucleic Acids Res. 0:0-0(0).
DR EMBL; X60213; CAA42770.1; -.
DR InterPro: IPR000349; Hepadnavir_surfa9.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 366 AA; 41057 MW; FDF3616EBC39629D CRC64;
```

| | | | | |
|-----------------------|--|---|----------------|-------------|
| Query Match | 89.5%; | Score 779.5; | DB 12; | Length 366; |
| Best Local Similarity | 88.3%; | Pred. No. 2.2e-64; | | |
| Matches 143; | Conservative | 10; | Mismatches 8; | Indels 1; |
| Gaps | 1; | | | |
| QY | 1 | MGQHPAKSMDVRRTEGGEILLNOLAGRMIPKGTLTWGSKFPTLDHVDHVTMEINTLQ | 60 | |
| Db | 37 | MGQHPAKSMDVRRTEGGEILLNOLAGRMIPKGTLTWGSKFPTLDHVDHVTMEINTLQ | 96 | |
| QY | 61 | NOGAWPAGARRVGLSNPTPOEIPQOWTPEEDQKAEAFRRYQERPPETTTIPPSPP | 120 | |
| Db | 97 | QOGAWPAGARRVGLSNPTPOEIPQOWTPEEDQKAEAFRRYQERPPETTTIAPTST | 156 | |
| QY | 121 | QWKLPQGDPLNGSLLETHPLYQS-EPAPVPVTKTPPLKKK | 161 | |
| Db | 157 | PKWKLPQGDPLNGSLLETHPLYQNPEAPVPVTKTPPLKKK | 198 | |
| RESULT | 6 | | | |
| Q9WFB3 | | PRELIMINARY; | PRT; | 329 AA. |
| ID | Q9WFB3 | | | |
| AC | Q9WFB3 | | | |
| DT | 01-NOV-1999 | (Tremblrel. 12, Created) | | |
| DT | 01-NOV-1999 | (Tremblrel. 12, Last sequence update) | | |
| DT | 01-JUN-2001 | (Tremblrel. 17, Last annotation update) | | |
| DE | PRES ANTIGEN. | | | |
| OS | snow goose hepatitis B virus. | | | |
| OC | Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus. | | | |
| OX | NCBI_TaxID=89623; | | | |
| RN | [1] | | | |
| RC | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=SGHBV1-7; | | | |
| RA | Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K., | | | |
| RA | Will H.; | | | |
| RT | "Identification and functional analysis of a new avian hepatitis B | | | |
| RL | virus naturally infecting snow geese (Anser caerulescens)."; | | | |
| DR | Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF110999; AAD21997.1; | | | |
| DR | InterPro; IPR000349; Hepadnavir_surfa. | | | |
| DR | Pfam; PF00695; VMSA; 2. | | | |
| SQ | SEQUENCE 329 AA; 36602 MW; 46DAB552978B0F27 CRC64; | | | |
| Query Match | 83.1%; | Score 724; | DB 12; | Length 329; |
| Best Local Similarity | 85.2%; | Pred. No. 2.6e-59; | | |
| Matches 138; | Conservative | 8; | Mismatches 14; | Indels 2; |
| Gaps | 2; | | | |
| QY | 1 | MGQHPAKSMDVRRTEGGEILLNOLAGRMIPKGTLTWGSKFPTLDHVDHVTMEINTLQ | 60 | |
| Db | 1 | MGQHPAKSMDVRRTEGGEILLNOLAGRMIPKGTLTWGSKFPTLDHVDHVTMEINTLQ | 60 | |
| QY | 61 | NOGAWPAGARRVGLSNPTPOEIPQOWTPEEDQKAEAFRRYQERPPETTTIPPSPP | 120 | |
| Db | 61 | KQAWPEAGARRVGLSNPTPOEIPQOWTPEEDQKAEAFRRYQERPPETTTIPP-TPT | 119 | |
| QY | 121 | QWKLPQGDPLNGSLLETHPLYQ-SEPAPVPVTKTPPLKKK | 161 | |
| Db | 120 | PKWKLPQGDPLNGSLLETHPLYQNPEAPVPVTKTPPLKKK | 161 | |
| RESULT | 7 | | | |
| Q9WFA9 | | PRELIMINARY; | PRT; | 329 AA. |
| ID | Q9WFA9 | | | |
| AC | Q9WFA9; | | | |
| DT | 01-NOV-1999 | (Tremblrel. 12, Created) | | |
| DT | 01-NOV-1999 | (Tremblrel. 12, Last sequence update) | | |
| DT | 01-JUN-2001 | (Tremblrel. 17, Last annotation update) | | |
| DE | PRES ANTIGEN. | | | |
| OS | snow goose hepatitis B virus. | | | |
| OC | Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus. | | | |
| OX | NCBI_TaxID=89623; | | | |
| RN | [1] | | | |

```
Q9WFB6
ID Q9WFB6 PRELIMINARY; PRT; 329 AA.
AC Q9WFB6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PRES ANTIGEN.
OS snow goose hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=89623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SGBV1-9;
RA Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K.,
RA Will H.;
RT "Identification and functional analysis of a new avian hepatitis B
RT virus naturally infecting snow geese (Anser caerulescens).";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF111000; AAD22002.1; -
DR InterPro: IPR000349; Hepadnavir_surfa.
DR Pfam: PF00695; vmsa; 2.
SQ SEQUENCE 329 AA; 36486 MW; 6A85C3EEEF3FA526 CRC64;

Query Match 82.3%; Score 717; DB 12; Length 329;
Best Local Similarity 84.6%; Pred. No. 1.2e-58;
Matches 137; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

QY 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTLQ 60

QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 KQAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 119

QY 121 QWKLQPGDDPLLGNSLLETHPLQ-SEPAVPVIKTPPLKK 161
Db 120 PKWLQPGDDPLLGKSLLETRLQTONSEPAVPVIKYPLVKK 161

RESULT 10
Q9WFA6
ID Q9WFA6 PRELIMINARY; PRT; 329 AA.
AC Q9WFA6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PRES ANTIGEN.
OS snow goose hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=89623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SGBV1-15;
RA Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K.,
RA Will H.;
RT "Identification and functional analysis of a new avian hepatitis B
RT virus naturally infecting snow geese (Anser caerulescens).";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF110997; AAD21987.1; -
DR InterPro: IPR000349; Hepadnavir_surfa.
DR Pfam: PF00695; vmsa; 2.
SQ SEQUENCE 329 AA; 36586 MW; 7C1928C4C0E87466 CRC64;

Query Match 82.0%; Score 714; DB 12; Length 329;
Best Local Similarity 84.0%; Pred. No. 2.2e-58;
Matches 136; Conservative 9; Mismatches 15; Indels 2; Gaps 2;

QY 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTLQ 60
```

```
QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 KQAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 119

QY 121 QWKLQPGDDPLLGNSLLETHPLQ-SEPAVPVIKTPPLKK 161
Db 120 PKWLQPGDDPLLGKSLLETRLQTONSEPAVPVIKYPLVKK 161

RESULT 11
Q67852
ID Q67852 PRELIMINARY; PRT; 327 AA.
AC Q67852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SURFACE PROTEIN.
GN PRES.
OS duck hepatitis B virus (DHBV).
OC Viruses; Retroviral viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi H., Cullen J.M., Newbold J.E.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M95589; AAA45749.1; -
DR InterPro: IPR000349; Hepadnavir_surfa.
DR Pfam: PF00695; vmsa; 2.
SQ SEQUENCE 327 AA; 36356 MW; 6875E959746DADEB CRC64;

Query Match 69.2%; Score 602.5; DB 12; Length 327;
Best Local Similarity 68.3%; Pred. No. 4.7e-48;
Matches 114; Conservative 17; Mismatches 23; Indels 13; Gaps 3;

QY 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVDHVTMEINTMQ 60

QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 KQAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 120

QY 121 --QWKLQPGDDPLLGNSLLETHPLQ----SEPAVPVIKTPPLKK 161
Db 121 KQWELKPGD-----PLLSTQPLRPAEPAEPDIPVITKPKVPKK 160

RESULT 12
Q9A4Y7
ID Q9A4Y7 PRELIMINARY; PRT; 418 AA.
AC Q9A4Y7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CONJUGAL TRANSFER PROTEIN TRBI.
GN CC2685.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Haft D.H.,
RA Shonay J.F., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
```

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005935; AAK24651.1; -
DR TIGR: CC2685; -
KW Complete proteome.
SQ SEQUENCE 418 AA; 44093 MW; 9AC05F75943954D0 CRC64;

Query Match 11.3%; Score 98.5; DB 2; Length 418;
Best Local Similarity 25.4%; Pred. No. 0.29;
Matches 44; Conservative 15; Mismatches 49; Indels 65; Gaps 8;
QY 34 LTW-----SGKFTPLDHLVDHVQTMEEINTLQN---QGAWPAGAG----- 70
DB 56 LAWFALSPRALQAAPKAPTVSSAVDRGTSAAEAVRQLPSYAQGGAPPRGLPPLPGDLGRAV 115
QY 71 ----RRVGL-----SNPTPQETPQP-----QWTEEDOKAREAFRRYQER--- 107
DB 116 INHQRRDGLVGDAGSDTP--MPQSTDOQAQVDAERQWREQAQARAGYVWQVTRRAE 173
QY 108 -----PPETTTIPSSPPQWKLPQGD-----PLIGNQSLLETHPL 143
DB 174 AAVAPNLLASPEFATPPVPSSPQIGEPGDQNRKMAFGAVPGGESIYTAHQL 226

RESULT 13
ID Q9Y4F2 PRELIMINARY; PRT; 1172 AA.
AC Q9Y4F2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE KIAA0442 PROTEIN (FRAGMENT).
GN KIAA0442.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
DR EMBL: AB007902; BAA23714.1; -
DR InterPro: IPR002965; P_Rich_extensn.
DR PRINTS: PR01217; PRICEXTENSN.
FT NON_TER 1
SQ SEQUENCE 1172 AA; 129749 MW; DCDD5B59347AFC29 CRC64;

Query Match 11.3%; Score 98; DB 4; Length 1172;
Best Local Similarity 29.6%; Pred. No. 1;
Matches 29; Conservative 13; Mismatches 42; Indels 14; Gaps 3;
QY 73 VGLSNPTP---QEIQQPQWTEEDQKARE---AFRRYQERPEPTTTIPSSPPQWKLPQ 126
DB 206 VVLKDCFOVAQPIQIQQTETQLRAPSPDPLVQRTAPPQPPPLUSTQPPQGPPEAQLQ 265
QY 127 GDDPLL-----GNQSLLETHPLHYQSEPAVPVIKTP 156
DB 266 APQVQVQRPQRPQSTQLLHQLNLPVQAHPSAQSLSQP 303
RESULT 14
ID Q9BTI7 PRELIMINARY; PRT; 548 AA.
AC Q9BTI7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR IMAGE:3453235) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC003681; AAH03681.1; -
FT NON_TER 1
SQ SEQUENCE 548 AA; 59695 MW; B156923EAAC01DA9 CRC64;

Query Match 11.0%; Score 96; DB 4; Length 548;
Best Local Similarity 22.4%; Pred. No. 0.68;
Matches 37; Conservative 21; Mismatches 47; Indels 60; Gaps 8;
QY 5 PAKSNDVRRRIEGGEILLNQLAGRMIPKGTLTWSGKFPTLDHVLVDHVQTMEEINTLQNOGA 64
DB 230 PAKPQPKMKTSGPV---MGGGLPP-----PPIKHMD-----IGTWDNKGK 268
QY 65 WPAGAGRRVGLSNPTPQEIPOPTPEEDQKARE-----AFRRYQERPEPTTTIP 115
DB 269 VPKA-----PVPQAQSPQAAPQQAQPLPAQPPALAPQQAQSPQQAQPPQTRRWA 319
QY 116 P-----SSPPQWKLPQGDPLIGNQSLLETHPLQ 145
DB 320 PRNRNAAFQSGGAGSDSNSPGNVQPNAP-----SVESHPVLE 358

RESULT 15
ID Q9BYJ9 PRELIMINARY; PRT; 559 AA.
AC Q9BYJ9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DJ963E22.1 (NOVEL PROTEIN SIMILAR TO NY-REN-2 ANTIGEN) (FRAGMENT).
GN DJ963E22.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL096828; CAC09391.3; -
FT NON_TER 1
SQ SEQUENCE 559 AA; 60873 MW; DB8AA15636130E18 CRC64;

Query Match 11.0%; Score 96; DB 4; Length 559;
Best Local Similarity 22.4%; Pred. No. 0.69;
Matches 37; Conservative 21; Mismatches 47; Indels 60; Gaps 8;
QY 5 PAKSNDVRRRIEGGEILLNQLAGRMIPKGTLTWSGKFPTLDHVLVDHVQTMEEINTLQNOGA 64
DB 241 PAKPQPKMKTSGPV---MGGGLPP-----PPIKHMD-----IGTWDNKGK 279
QY 65 WPAGAGRRVGLSNPTPQEIPOPTPEEDQKARE-----AFRRYQERPEPTTTIP 115
DB 280 VPKA-----PVPQAQSPQAAPQQAQPLPAQPPALAPQQAQSPQQAQPPQTRRWA 330
QY 116 P-----SSPPQWKLPQGDPLIGNQSLLETHPLQ 145
DB 331 PRNRNAAFQSGGAGSDSNSPGNVQPNAP-----SVESHPVLE 369

Search completed: May 13, 2002, 11:23:19
Job time: 196 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:20:23 ; Search time 11.77 Seconds
(without alignments)
501.532 Million cell updates/sec

Title: US-09-818-066-34

Perfect score: 871

Sequence: 1 MGQHPAKSMDVRRIEGEL.....PLYQSEPAVPVTKPLPKKK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Match | Length | DB | ID | Description |
|------------|-------|---------|-------|--------|------------|--------|--------------|
| 1 | 871 | 100.0 | 328 | 1 | VNSA_HPBDU | P03145 | duck hepati |
| 2 | 793.5 | 91.1 | 365 | 1 | VNSA_HPBDU | P30029 | duck hepati |
| 3 | 782.5 | 89.8 | 366 | 1 | VNSA_HPBDU | P17195 | duck hepati |
| 4 | 777.5 | 89.3 | 366 | 1 | VNSA_HPBDU | P17194 | duck hepati |
| 5 | 411.5 | 47.2 | 335 | 1 | VNSA_HPBDU | P13847 | heron hepati |
| 6 | 102 | 11.7 | 3530 | 1 | MY15_HUMAN | Q9ukn7 | homo sapien |
| 7 | 92 | 10.6 | 746 | 1 | ABL_MLVAB | P00521 | abelson mur |
| 8 | 92 | 10.6 | 1123 | 1 | ABL_MOUSE | P00520 | mus musculus |
| 9 | 89 | 10.2 | 654 | 1 | ICAL_RAT | P27321 | rattus norv |
| 10 | 89 | 10.2 | 1319 | 1 | MN1_HUMAN | Q10571 | homo sapien |
| 11 | 88.5 | 10.2 | 1194 | 1 | MGR1_HUMAN | Q13255 | homo sapien |
| 12 | 86 | 9.9 | 704 | 1 | MYBB_MOUSE | P48972 | mus musculus |
| 13 | 86 | 9.9 | 1500 | 1 | SSP5_STRGN | P16952 | stretococc |
| 14 | 85.5 | 9.8 | 938 | 1 | EBN4_EBV | P03203 | epstein-bar |
| 15 | 85 | 9.8 | 514 | 1 | MEFD_MOUSE | Q63943 | mus musculus |
| 16 | 85 | 9.8 | 565 | 1 | MOT8_MOUSE | O70324 | mus musculus |
| 17 | 84.5 | 9.7 | 510 | 1 | ERM_HUMAN | P41161 | homo sapien |
| 18 | 84.5 | 9.7 | 805 | 1 | YGM6_YEAST | P53086 | saccharomyc |
| 19 | 84 | 9.6 | 417 | 1 | OPD2_BACSU | P16263 | bacillus su |
| 20 | 84 | 9.6 | 708 | 1 | VP40_HCVFA | P16753 | human cytom |
| 21 | 84 | 9.6 | 1902 | 1 | SMF1_HUMAN | O14497 | homo sapien |
| 22 | 83 | 9.5 | 407 | 1 | ARRH_LOCM1 | P32122 | locusta mig |
| 23 | 82.5 | 9.5 | 757 | 1 | HT16_HYDAT | P53356 | hydra atten |
| 24 | 82.5 | 9.5 | 1257 | 1 | PGCN_RAT | P55067 | glycine max |
| 25 | 82 | 9.4 | 309 | 1 | NO75_SOYBN | P08297 | glycine max |
| 26 | 82 | 9.4 | 1664 | 1 | INT1_CANAL | P33705 | candida alb |
| 27 | 82 | 9.4 | 5376 | 1 | ZAN_MOUSE | O88799 | mus musculus |
| 28 | 81.5 | 9.4 | 1199 | 1 | MGR1_RAT | P23385 | rattus norv |
| 29 | 81 | 9.3 | 2469 | 1 | TEGU_HSVSA | Q01056 | herpesvirus |
| 30 | 80.5 | 9.2 | 269 | 1 | YA53_ARCFU | O29209 | archaeoglob |
| 31 | 80.5 | 9.2 | 484 | 1 | LPG4_VICFA | P05190 | vicia faba |
| 32 | 80.5 | 9.2 | 1001 | 1 | IF2_SYNY3 | P2689 | synechocyst |
| 33 | 80 | 9.2 | 296 | 1 | GDA6_WHEAT | P04726 | trititum ae |

RESULT 1
VNSA_HPBDU

ID VNSA_HPBDU STANDARD: PRT; 328 AA.

AC P03145;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE MAJOR SURFACE ANTIGEN PRECURSOR.

GN S.

OS Duck hepatitis B virus (DHBV).

OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.

OX NCBI_TaxID=12639;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84138772; PubMed=6699938;

RA Mandart E., Kay A., Galibert F.;

RT "Nucleotide sequence of a cloned duck hepatitis B virus genome:

RT comparison with woodchuck and human hepatitis B virus sequences.";

RL J. Virol. 49:782-792(1984).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-ISOLATE DHBV F1-6;

RX MEDLINE=91045092; PubMed=2235507;

RA Mattes F., Tong S., Teubner K., Blum H.E.;

RT "Complete nucleotide sequence of a German duck hepatitis B virus.";

RN Nucleic Acids Res. 18:6140-6140(1990).

RP MYRISTOYLATION.

RX MEDLINE=91135002; PubMed=1994583;

RA Macrae D.R., Brusa V., Ganem D.;

RT "Myristylation of a duck hepatitis B virus envelope protein is

essential for infectivity but not for virus assembly.";

RL Virology 181:359-363(1991).

CC -!- PTM: MYRISTOYLATION CONTRIBUTES IMPORTANTLY TO DHBV INFECTIVITY.

CC IT IS MOST LIKELY REQUIRED FOR AN EARLY STEP OF THE LIFE CYCLE

CC INVOLVING THE ENTRY OR UNCOATING OF VIRUS PARTICLES.

CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
EMBL; X12798; CAB57224.1; -
DR PIR; A03710; SAVLD.
DR PIR; S12845; S12845.
DR InterPro; IPR000349; Hepadnavir_surfac.
DR Pfam; PF00695; VNSA; 2.
KW Antigen; Myristate; Envelope protein; Lipoprotein.
FT PROPEP 1 161
FT CHAIN 162 328 MAJOR SURFACE ANTIGEN.
FT LIPID 2 2 MYRISTATE.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT

SQ SEQUENCE 328 AA; 36230 MW; B2D771241E407456 CRC64;

Query Match 100.0%; Score 871; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 7.9e-64;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 60
|||||
DB 1 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 60
|||||
QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 161
|||||
DB 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 161
|||||

RESULT 2

VMSA_HPBDC STANDARD; PRT; 365 AA.
ID VMSA_HPBDC
AC P30029;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (strain China) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=31510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91045091; PubMed=2235506;
RA Tong S., Mattes F., Teubner K., Blum H.E.;
RT "Complete nucleotide sequence of a Chinese duck hepatitis B virus.";
RL Nucleic Acids Res. 18:6139-6139(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21953; AAA45746.1; -
DR PIR; S12842; SAVLWE.
DR InterPro; IPR000349; Hepadnavir_surfa.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 198
FT CHAIN 199 365 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 365 AA; 40511 MW; 67F3A4174CB7D884 CRC64;

Query Match 91.1%; Score 793.5; DB 1; Length 365;
Best Local Similarity 90.7%; Pred. No. 1.7e-57;
Matches 147; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 60
|||||
DB 37 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 96
|||||
QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 97 QGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 156
|||||
QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 161
|||||
DB 157 PKLQPGDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 198
|||||

RESULT 3

VMSA_HPBDC STANDARD; PRT; 366 AA.
ID VMSA_HPBDC
AC P17195;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (white Shanghai duck isolate S31) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=10440;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085807; PubMed=2596031;
RA Uchida M., Esumi M., Shikata T.;
RT "Molecular cloning and sequence analysis of duck hepatitis B virus
RT genomes of a new variant isolated from Shanghai ducks.";
RL Virology 173:600-606(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32991; AAA45752.1; ALT_INIT.
DR PIR; D33746; SAVLWD.
DR InterPro; IPR000349; Hepadnavir_surfa.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 199
FT CHAIN 200 366 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 366 AA; 40858 MW; CF60E78B7B2FCD52 CRC64;

Query Match 89.8%; Score 782.5; DB 1; Length 366;
Best Local Similarity 88.3%; Pred. No. 1.3e-56;
Matches 143; Conservative 11; Mismatches 7; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 60
|||||
DB 37 MGQHPAKSMVRRIEGELLNLAGRMIPKGLTWSGKFPDLDHVDHVQTMEEINTLQ 96
|||||
QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 120
|||||
DB 97 QGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSSPP 156
|||||
QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 161
|||||
DB 157 PKLQPGDDPDLGNQSLLETHPLYSQSEPAVPVKTPLPKK 198
|||||

RESULT 4

VMSA_HPBDC STANDARD; PRT; 366 AA.
ID VMSA_HPBDC
AC P17194;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (brown Shanghai duck isolate S5) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=10439;
RN [1]
RP SEQUENCE FROM N.A.


```
RX MEDLINE-90085807; PubMed=2596031;
RA Uchida M., Esumi M., Shikata T.;
RT "Molecular cloning and sequence analysis of duck hepatitis B virus
RL genomes of a new variant isolated from Shanghai ducks.";
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32990; AAA45755.1; ALT_INIT.
DR PIR; C33746; SAVLBD.
DR InterPro; IPR000349; Hepadnavir_surfaG.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 199
FT CHAIN 200 366 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 170 366 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 366 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 366 AA; 40897 MW; 5B72879A182EFF38 CRC64;

Query Match 89.3%; Score 777.5; DB 1; Length 366;
Best Local Similarity 87.7%; Pred. No. 3.4e-56;
Matches 142; Conservative 11; Mismatches 8; Indels 1; Gaps 1;

Qy 1 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEINPLQ 60
Db 1 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEINPLQ 60
Qy 37 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEINPLQ 96
Db 37 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEINPLQ 96
Qy 61 NOGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPPSGPP 120
Db 61 NOGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPPSGPP 120
Qy 97 QGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPPSGPP 156
Db 97 QGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPPSGPP 156
Qy 121 QWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 161
Db 121 QWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 161
Qy 157 PWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 198
Db 157 PWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 198

RESULT 5
VMSA_HPBHE STANDARD; PRT; 335 AA.
AC P13847; 067853;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Heron hepatitis b virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=28300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88333160; PubMed=3418788;
RA Spengel R., Kaleta E.F., Will H.;
RT "Isolation and characterization of a hepatitis B virus endemic in
RT herons.";
RL J. Virol. 62:3832-3839(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22056; AAA45739.1;
DR EMBL; M22056; AAA45740.1; ALT_INIT.
```

```
DR PIR; B30082; SAVLHH.
DR InterPro; IPR000349; Hepadnavir_surfaG.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 166
FT CHAIN 167 335 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 265 335 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 335 AA; 37218 MW; E90F7DA6A0623354 CRC64;

Query Match 47.2%; Score 411.5; DB 1; Length 335;
Best Local Similarity 48.5%; Pred. No. 1.5e-26;
Matches 83; Conservative 23; Mismatches 48; Indels 17; Gaps 5;

Qy 1 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEIN 57
Db 1 MGQHPAKSMQVRRTEGGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVDVOTMEIN 57
Qy 58 TLQNOGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPP- 116
Db 58 TLQNOGAWPAGAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPP- 116
Qy 61 TLQAGGHWPGEATARRRLGLDQPRTPP-PPTWTTEEDKKAKEFFKQYQENRRPKPAETAPP 119
Db 61 TLQAGGHWPGEATARRRLGLDQPRTPP-PPTWTTEEDKKAKEFFKQYQENRRPKPAETAPP 119
Qy 117 -----SPPQWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 161
Db 117 -----SPPQWKLQPGDDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK 161
Qy 120 ITELHAAEPQWPKISP-EDPLLKAKALIPV-----KEPEVPILKVKPLTKN 164
Db 120 ITELHAAEPQWPKISP-EDPLLKAKALIPV-----KEPEVPILKVKPLTKN 164

RESULT 6
MY15_HUMAN STANDARD; PRT; 3530 AA.
ID MY15_HUMAN
AC Q90KN7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN XV (UNCONVENTIONAL MYOSIN-15).
GN MYO15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE-20021762; PubMed=10552926;
RA Liang Y., Wang A., Belyantseva I.A., Anderson D.W., Probst F.J.,
RA Barber T.D., Miller W., Touchman J.W., Jin L., Sullivan S.L.,
RA Sellers J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman T.B.,
RA Fridell R.A.;
RT "Characterization of the human and mouse unconventional myosin XV
RT genes responsible for hereditary deafness DFNB3 and shaker 2.";
RL Genomics 61:243-258(1999).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND VARIANTS DFNB3 TYR-2111 AND PHE-2113.
RX MEDLINE-98267311; PubMed=9603736;
RA Wang A., Liang Y., Fridell R.A., Probst F.J., Wilcox E.R.,
RA Touchman J.W., Morton C.C., Morell R.J., Noben-Trauth K., Camper S.A.,
RA Friedman T.B.;
RT "Association of unconventional myosin MYO15 mutations with human
RT nonsyndromic deafness DFNB3.";
RL Science 280:1447-1451(1998).
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
CC SIMILARITY). MAY PLAY A ROLE IN THE FORMATION OR MAINTENANCE OF
CC THE ACTIN-RICH STRUCTURES OF THE INNER EAR SENSORY HAIR CELLS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PITUITARY. ALSO EXPRESSED
CC AT LOWER LEVELS IN ADULT BRAIN, KIDNEY, LIVER, LUNG, PANCREAS,
CC PLACENTA AND SKELETAL MUSCLE. NOT EXPRESSED IN BRAIN. IN THE
CC PITUITARY, HIGHLY EXPRESSED IN ANTERIOR GLAND CELLS.
CC -1- DISEASE: DEFECTS IN MYO15 ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
CC FORM OF NONSYNDROMIC DEAFNESS (DFNB3).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
```

CC -!- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 IQ DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF144094; AAF05903.1; -
DR EMBL: AF051976; -; NOT_ANNOTATED_CDS.
DR HSSP: P08799; 1MND.
DR MIM: 602666; -
DR MIM: 600316; -
DR InterPro: IPR000299; Band_4.1.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR000857; MYTH4.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00963; myosin_head; 1.
DR Pfam: PF00784; MYTH4; 2.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MYSC; 1.
DR SMART: SM00139; MYTH4; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00660; BAND_4.1_1; FALSE_NEG.
DR PROSITE: PS00661; BAND_4.1_2; FALSE_NEG.
DR PROSITE: PS50057; BAND_4.1_3; 1.
DR PROSITE: PS50096; IQ; 3.
DR PROSITE: PS50002; SH3; FALSE_NEG.
KW Myosin; ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
KW Calmodulin-binding; Disease domain.
FT DOMAIN 1 1887 HEAD OR MOTOR DOMAIN.
FT DOMAIN 2030 3530 NECK OR REGULATORY DOMAIN.
FT DOMAIN 1888 2029 TAIL.
FT DOMAIN 1323 1350 COILED COIL (POTENTIAL).
FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
FT DOMAIN 1902 1924 IQ 1.
FT DOMAIN 1925 1954 IQ 2.
FT DOMAIN 1955 1976 IQ 3.
FT DOMAIN 2867 2953 SH3.
FT DOMAIN 3206 3443 BAND 4.1-LIKE.
FT NP_BIND 1315 1322 ATP (POTENTIAL).
FT VARIANT 2111 2111 N -> Y (IN DFNB3; FAMILY FROM BENGKALA).
FT VARIANT 2113 2113 I -> F (IN DFNB3; INDIAN FAMILY).
FT FTID-VAR_010303.
FT FTID-VAR_010304.
SQ SEQUENCE 3530 AA; 395171 MW; 3D103923D4BCBEA4 CRC64;

Query Match 11.7%; Score 102; DB 1; Length 3530;
Best Local Similarity 31.1%; Pred. No. 2.7;
Matches 33; Conservative 9; Mismatches 44; Indels 20; Gaps 5;

QY 65 WPAGAGRRVCLSNPTP-----QEIPQOWTEED--QKAREAFRRYQEEPPETT 112
Db 943 WPGAGSRRGFRPPVPENPFQLGLGVPSFTLQEDPADMTVRFLGRHHEPGGQLT 1002
QY 113 TIPPSSPPQWKLQGDPLGNLSL-LETHPLYQSEPAVPVKTTP 157
Db 1003 KSAGPIPE-----KPEEATLGDPQLPAETKP---PTPAPKDVTPP 1041

RESULT 7
ID ABL_MLVAB STANDARD; PRT; 746 AA.
AC P00521;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE TYROSINE-PROTEIN KINASE TRANSFORMING PROTEIN ABL (BC 2.7.1.112).
GN V-ABL.
OS Abelson murine leukemia virus.
OC Viruses; Retrovirdae; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxId=11788;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83221648; PubMed=6304726;
RA Reddy E.P., Smith M.J., Srinivasan A.;
RT "Nucleotide sequence of Abelson murine leukemia virus genome:
RT structural similarity of its transforming gene product to other onc
RT gene products with tyrosine-specific kinase activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
RN [2]
RP SEQUENCE OF 233-327 FROM N.A.
RX MEDLINE=83245023; PubMed=6191223;
RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
RT "Homology between phosphotyrosine acceptor site of human c-abl and
RT viral oncogene products.";
RL Nature 304:167-169(1983).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
CC POLYPEPTIDE.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE ABL SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01541; -; NOT_ANNOTATED_CDS.
DR EMBL: K00010; AAA46470.1; -
DR PIR: A00627; TVMVGM.
DR HSSP: P00519; 2ABL.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000980; SH2_
DR InterPro: IPR001245; Tyr_kin.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRINTS: PR00401; SH2DOMAIN.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TyrKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
KW Polyprotein; Transforming protein; Tyrosine-protein kinase;
KW Transferase; Oncogene; SH2 domain.
FT DOMAIN 13 103 SH2.
FT DOMAIN 128 379 PROTEIN KINASE.
SQ SEQUENCE 746 AA; 81872 MW; B9072FFF55FE9257 CRC64;

Query Match 10.6%; Score 92; DB 1; Length 746;
Best Local Similarity 29.0%; Pred. No. 2.9;
Matches 31; Conservative 12; Mismatches 38; Indels 26; Gaps 5;

QY 53 MEEINTLQNOGAWPAGARRVGLSNPTPQEIPOQWTEEDQKAREAFRRYQEEPPETT 112
Db 1003 KSAGPIPE-----KPEEATLGDPQLPAETKP---PTPAPKDVTPP 1041

[illegible]

(correctly constructed) and the other


```
FT TRANSMEM 1475 1495 POTENTIAL.
FT DOMAIN 1496 1500 CYTOPLASMIC (POTENTIAL).
FT SIMILAR 164 470 TO M PROTEIN OF S.PYOGENES.
FT DOMAIN 164 470 4 X APPROXIMATE TANDEM REPEATS, HRI.
FT REPEAT 164 241 1.
FT REPEAT 242 323 2.
FT REPEAT 324 405 3.
FT REPEAT 406 470 4.
FT DOMAIN 771 887 3 X APPROXIMATE TANDEM REPEATS, PRI.
FT DOMAIN 1414 1436 PRO-RICH (PR2).
FT DOMAIN 1467 1472 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT CA-BIND 220 235 POTENTIAL.
FT CA-BIND 301 316 POTENTIAL.
FT CA-BIND 931 950 POTENTIAL.
FT CA-BIND 1300 1315 POTENTIAL.
SQ SEQUENCE 1500 AA; 164552 MW; DCF190E7D44D889F CRC64;

Query Match 9.9%; Score 86; DB 1; Length 1500;
Best Local Similarity 23.7%; Pred. No. 20;
Matches 27; Conservative 18; Mismatches 33; Indels 36; Gaps 4;

QY 76 SNPTPOEIPQOWTPEEDOKAREAFRRYOERPPETTTIPPS-----SPQWKQLQ 125
DQ 788 NEPTP-----PVKTPDQPEPSKPEEPTYTEKPLEPAPVAPSYENEPTPPVKTPDQPEPS 842
QY 126 PGDDPLLNQSLLETHPL---YOSPPAVPV-----IKTPPL 158
DQ 843 KPEEPNYETKPLEPAPVAPSYENEPTPPVKIPDQPEPSKPEEPTYPDPLTPPL 896

RESULT 14
EBNA_EBV
ID EBNA_EBV STANDARD; PRT; 938 AA.
AC P03203;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE EBNA-4 NUCLEAR PROTEIN (EBNA-3B).
GN BRF2A-BRF2B.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrall B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=90204695; PubMed=2157061;
RA Kerdiles B., Walls H., Perricaudet M., Joab I.;
RT "cDNA cloning and transient expression of the Epstein-Barr virus-
RT determined nuclear antigen EBNA3B in human cells and identification
RT of novel transcripts from its coding region.";
RL J. Virol. 64:1812-1816(1990).
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=90266473; PubMed=2161150;
RA Pettit L., Sample C., Kieff E.;
RT "Subnuclear localization and phosphorylation of Epstein-Barr virus
RT latent infection nuclear proteins.";
RL Virology 176:563-574(1990).
CC -!- FUNCTION: INVOLVED IN LATENT CYCLE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, 5, AND 6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC -----
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01555; CAA24858.1; -
CC PIR; A03765; Q0BE24.
CC Nuclear protein.
SQ SEQUENCE 938 AA; 102860 MW; 98A513AAFC0676FD CRC64;

Query Match 9.8%; Score 85.5; DB 1; Length 938;
Best Local Similarity 23.2%; Pred. No. 13;
Matches 36; Conservative 21; Mismatches 53; Indels 45; Gaps 7;

QY 25 AGRMIPKGTLTWSGKF-----PTLDHVL-----HYVTMEEINTLQNOGAMP 66
DQ 534 AGR---RGPCVFTGDLGIESDEPASTEVPVHDQLLPAGDPLEIQTLSPTTSQLSSAP 590
QY 67 AGAGRRVGLSNPTQOEIPQOWTPEEDOKAREAFRRYOERPPETTTIPPSPPQWKLOP 126
DQ 591 SCA-----QTPWVVPQSQTPDDPTK-----QSRPPET-----AAPROMPL 628
QY 127 GDDPL--LGNQSLLETHPLYSQSEPAVPVVKTPPLK 159
DQ 629 RPIPMRLMQPIPFNHPVGPVTPHOTPOVEITPYK 663

RESULT 15
MEFD_MOUSE
ID MEFD_MOUSE STANDARD; PRT; 514 AA.
AC Q63943; Q63944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOCYTE-SPECIFIC ENHANCER FACTOR 2D.
GN MEF2D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94158837; PubMed=8114702;
RA Martin J.F., Miano J.M., Hustad C.M., Copeland N.G., Jenkins N.A.,
RA Olson E.N.;
RT "A Mef2 gene that generates a muscle-specific isoform via alternative
RT mRNA splicing.";
RL Mol. Cell. Biol. 14:1647-1656(1994).
CC -!- FUNCTION: TRANSCRIPTION FACTOR WHICH BINDS SPECIFICALLY TO THE
CC MEF2 ELEMENT PRESENT IN THE REGULATORY REGIONS OF MANY MUSCLE-
CC SPECIFIC GENES. ACTIVATES TRANSCRIPTION VIA THIS ELEMENT. MAY BE
CC INVOLVED IN MUSCLE-SPECIFIC AND/OR GROWTH FACTOR-RELATED
CC TRANSCRIPTION.
CC -!- SUBUNIT: HETERODIMER.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; NON-MUSCLE AND MUSCLE; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: ONE OF THE TWO SPLICED FORMS IS MUSCLE-
CC SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS. MEF2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

DR EMBL; S68893; AAB29973.1; -;
DR EMBL; S68895; AAB29974.1; -;
DR HSP; P11831; ISRS.
DR MGD; MGI:99533; Mef2d.
DR InterPro; IPR002100; MADS-box.
DR Pfam; PF00319; SRF-TF; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS00350; MADS_BOX_1; 1.
DR PROSITE; PS00350; MADS_BOX_2; 1.
KW Transcription regulation; Nuclear protein; DNA-binding; Activator;
KW Multigene family; Alternative splicing.
FT DOMAIN 3 57 MADS.
FT DNA_BIND 58 86 MEF2-TYPE (POTENTIAL).
FT DOMAIN 252 255 POLY-PRO.
FT DOMAIN 444 449 POLY-PRO.
FT DOMAIN 3 31 ARG/LYS-RICH (BASIC).
FT DOMAIN 365 402 GLN/PRO-RICH.
FT VARSPLIC 87 132 TLKKGNGCDSPEDGDSLEQSPLEDKYRASELDGL
FRYVG -> ALHNDRECSPEVDEAFALTPTQTEKKYKID
EEYKKIDEEDFKMOSYRLA (IN MUSCLE
ISOFORM).
FT VARSPLIC 286 292 MISSING (IN MUSCLE ISOFORM).
SQ SEQUENCE 514 AA; 54992 MW; 1D25A80DB3E04F43 CRC64;

Query Match 9.88; Score 85; DB 1; Length 514;
Best Local Similarity 23.3%; Pred. No. 7.1;
Matches 47; Conservative 27; Mismatches 58; Indels 70; Gaps 11;
QY 5 PAKSMDVRRIT--EGGETLLNLAG-----RMIPKG 32
DB 264 PSRKPDLRVITSGGKGLMHHLGDHLDLNNQRLGVSTHSLTTPVSVATPSSLQ 323
QY 33 TLWSGKFPPTLDVLDHVQTMEEINTLQNGAWPAGARRVGLSNPTTPQEIPOQTPEE 92
DB 324 -LPFS-SMPTA-VNTDYQLPSAELSLP---AFSSPAG--LALGNVTAWQQPQQPQP 375
QY 93 DQKAREAFRYQERPPETTIPSPSPQW-----KLPQDD-PLLGNSL 137
DB 376 PQ-----PPSQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQP 422
QY 138 LETHP--LYQSEPAVPVKTTP 157
DB 423 VTTHPHISIKSEPVSPRSRSP 444

Search completed: May 13, 2002, 11:23:37
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:19:13 ; Search time 14.4 Seconds
(without alignments)
851.674 Million cell updates/sec

Title: US-09-818-066-34
Perfect score: 871
Sequence: 1 MQGHPAKSMVRRIEGGEIL.....PLYQSEPAVPVIKTPPLKKK 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 871 | 100.0 | 364 | 1 SAVLWE | large surface anti |
| 2 | 793.5 | 91.1 | 365 | 1 SAVLWE | large surface anti |
| 3 | 782.5 | 89.8 | 366 | 1 SAVLWD | large surface anti |
| 4 | 777.5 | 89.3 | 366 | 1 SAVLBD | large surface anti |
| 5 | 692 | 79.4 | 158 | 2 JC1095 | Pre-S protein - du |
| 6 | 411.5 | 47.2 | 335 | 1 SAVLHH | large surface anti |
| 7 | 102 | 11.7 | 3530 | 2 A59266 | unconventional myo |
| 8 | 98 | 11.3 | 1172 | 2 T00065 | hypothetical prote |
| 9 | 95.5 | 11.0 | 633 | 2 T05005 | hypothetical prote |
| 10 | 92 | 10.6 | 539 | 2 T47405 | hypothetical prote |
| 11 | 92 | 10.6 | 981 | 1 FOMVGM | gag-abl polyprotel |
| 12 | 92 | 10.6 | 1123 | 2 A39362 | kinase-related tra |
| 13 | 91.5 | 10.5 | 299 | 2 T09792 | proline-rich prote |
| 14 | 90 | 10.3 | 407 | 2 C70816 | hypothetical prote |
| 15 | 89 | 10.2 | 283 | 2 T16348 | hypothetical prote |
| 16 | 89 | 10.2 | 603 | 2 S15074 | calpastatin - rat |
| 17 | 89 | 10.2 | 654 | 2 T10772 | calpastatin - rat |
| 18 | 87 | 10.0 | 296 | 2 A27319 | gliadin - wheat |
| 19 | 86.5 | 9.9 | 309 | 2 T29293 | hypothetical prote |
| 20 | 86.5 | 9.9 | 369 | 2 S20500 | hydroxyproline-ric |
| 21 | 86.5 | 9.9 | 564 | 2 S37241 | legumin B - fava b |
| 22 | 86.5 | 9.9 | 1952 | 2 T48814 | hypothetical prote |
| 23 | 86 | 9.9 | 285 | 2 A41826 | probable pheromone |
| 24 | 86 | 9.9 | 704 | 1 S33704 | transforming prote |
| 25 | 86 | 9.9 | 1473 | 2 A35186 | salivary agglutini |
| 26 | 85.5 | 9.8 | 430 | 2 JC2301 | hypothetical 47.8K |
| 27 | 85.5 | 9.8 | 938 | 1 Q0BE24 | nuclear antigen EB |
| 28 | 85 | 9.8 | 224 | 2 T03238 | extensin (clone Ex |
| 29 | 85 | 9.8 | 280 | 2 T03236 | extensin precursor |

30 85 9.8 377 2 A48018 mucin 7 precursor,
31 85 9.8 514 2 A56201 transcription fact
32 85 9.8 1001 2 T28897 hypothetical prote
33 84.5 9.7 359 2 T13478 hypothetical prote
34 84.5 9.7 510 1 S43692 transcription fact
35 84.5 9.7 805 2 S64238 kinesin-related pr
36 84 9.6 379 2 T05441 proline-rich prote
37 84 9.6 379 2 D85257 extensin-like prot
38 84 9.6 417 2 B32879 dihydroliipoamide S
39 84 9.6 443 2 T05540 hypothetical prote
40 84 9.6 708 1 Q0BE98 UL80 protein - hum
41 84 9.6 1142 2 T00022 B120 protein - hum
42 84 9.6 1274 2 T16251 hypothetical prote
43 83 9.5 407 2 A56607 arrestin homolog -
44 83 9.5 566 2 T06453 probable legumin B
45 83 9.5 875 2 T27495 hypothetical prote

ALIGNMENTS

RESULT 1

SAVLWD

large surface antigen - duck hepatitis virus

N:Contains: major surface antigen; middle surface antigen

C:Species: duck hepatitis virus, DHBV

C:Date: 20-Sep-1984 #sequence_revision 20-Sep-1984 #text_change 13-Mar-1997

C:Accession: A03710; S12845

R:Mandat: E.; Kay, A.; Galibert, F.

J. Virol. 49, 782-792, 1984

A:Title: Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison wi

A:Reference number: A92997; MUID:84138772

A:Accession: A03710

A:Molecule type: DNA

A:Residues: 1-364 <MAN>

A:Cross-references: CB:K01834

R:Mattes, F.; Tong, S.; Teubner, K.; Blum, H.E.

Nucleic Acids Res. 18, 6140, 1990

A:Title: Complete nucleotide sequence of a German duck hepatitis B virus.

A:Reference number: S12843; MUID:91045092

A:Accession: S12845

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 36-364 <MAT>

A:Cross-references: EMBL:X12798

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:89-364/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:198-364/Product: major surface antigen (gene S) #status predicted <MSA>

F:32,170,296/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 100.0%; Score 871; DB 1; Length 364;

Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQGHPAKSMVRRIEGGEILLNQLAGRIPKGTLTWSCKFPTLDHVLHVQTMEEINTLQ 60

Db 37 MQGHPAKSMVRRIEGGEILLNQLAGRIPKGTLTWSCKFPTLDHVLHVQTMEEINTLQ 96

Qy 61 NQAWPAGAGRRVGLSNPTPOEIPQOWTPEDQKARAFRRYOBERPETTTTPSSPP 120

Db 97 NQAWPAGAGRRVGLSNPTPOEIPQOWTPEDQKARAFRRYOBERPETTTTPSSPP 156

Qy 121 QWKLPQGGDPLLGNSLLETHPLYSQSEPAVPVIKTPPLKKK 161

Db 157 QWKLPQGGDPLLGNSLLETHPLYSQSEPAVPVIKTPPLKKK 197

RESULT 2

SAVLWE

[illegible]

QY 21 LNQLAGRMIPKCTLT--WSGKPTFLDVLHD-----VQTMEINTLQNGAWFA 67
Db 473 VQELLGK-----TGNVTSFW-----MDHIKTHCVSVPSVEEAAATREAVYNLQ-----WPP 519
QY 68 GAGRRY-----GLSNPTQOEIPQOWTPPEEDQKAREAFRRYQERPPETTTP 115
Db 520 NGRHLIAEFVRAEEVKELEAPLP---PQHQFOAQILSR-----PPPTALP 565
QY 116 PSSPPQWKLOQDDPLLGNQSLLETHPLXOSEPAVPVTKTTPPL 158
Db 566 P--PP-----PLAKPPHVVERLPLPPPPPIAPEQEPPPI 597
RESULT 10
T47405
Hypothetical protein F23N14.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47405
R:Blocker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24465
A:Accession: T47405
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <BLO>
A:Cross-references: EMBL:AL138638
A:Experimental source: cultivar Columbia; BAC clone F23N14
C:Genetics:
A:Map position: 3
A:Note: F23N14.70

Query Match 10.6%; Score 92; DB 2; Length 539;
Best Local Similarity 26.1%; Pred. No. 2.4;
Matches 37; Conservative 16; Mismatches 63; Indels 26; Gaps 7;
QY 29 IPKGTITWSGKF-PTLDHVLHDVQTMEEINTLQNGAWFAGAGRRV-----GLSNPT 79
Db 47 VPSSSATMASSFNPT--RILDHRASSHR---NRRGAFFASKRRRLVDEPIDYDLSNPA 100
QY 80 PQEIPQOWTPPEEDQKAREAFRRYQERPPETTTPPSPQWKLOQDDPLLGNQSLLE 139
Db 101 YQVLSITPLFA-----SGIGSIRELLSSPPPTTS---SQPPSVSIPPSPAPPLVLSDSKD 152
QY 140 THPLXOSEPAVPVTKTTPPLKKK 161
Db 153 AEPAGLTNPASAP----PSPLAPK 171

RESULT 11
F0MVGM
gag-abl polyprotein - Abelson murine leukemia virus
N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein p1
C:Species: Abelson murine leukemia virus
A:Note: host Mus sp. (mouse)
C:Date: 14-Nov-1983 #sequence_revision 09-Sep-1994 #text_change 11-Jun-1999
C:Accession: A03931; A00627; A93955
R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
Proc. Natl. Acad. Sci. U.S.A. 80, 3623-3627, 1983
A:Title: Nucleotide sequence of Abelson murine leukemia virus genome: structural similar
A:Reference number: A93955; MUID:83221648
A:Accession: A03931
A:Molecule type: DNA
A:Residues: 1-981 <RED>
A:Cross-references: GB:J02009; NID:g331887; PIDN:AAA46471.1; PID:g331888
A:Note: the authors translated the codon CGA for residue 186 as Glu
R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
Proc. Natl. Acad. Sci. U.S.A. 80, 7372, 1983
A:Reference number: A93980
A:Contents: annotation; erratum, residues 588-746
C:Genetics:
A:Gene: gag-abl

C:Superfamily: Abelson murine leukemia virus gag-abl polyprotein; protein kinase homo
C:Keywords: ATP; core protein; oncogene; phosphotransferase; polyprotein; transformin
F:1-131/Product: core protein p15 #status predicted <p15>
F:132-215/Product: inner coat protein p12 #status predicted <p12>
F:216-235/Region: amino end of core shell protein p30
F:248-338/Domain: SH2 homology <SH2>
F:361-621/Domain: protein kinase homology <KIN>
F:369-377/Region: protein kinase ATP-binding motif
F:392/Active site: Lys #status predicted

Query Match 10.6%; Score 92; DB 1; Length 981;
Best Local Similarity 29.0%; Pred. No. 5;
Matches 31; Conservative 12; Mismatches 38; Indels 26; Gaps 5;

QY 53 MBEINTLQNGAWFAGAGRRVGLSNPTQOEIPQOWTPPEEDQKAREAFRRYQERPPETT 112
Db 788 MDAVNTDPTK-AGPCGE---GLRKVPVPSVKPQST-----AKPPGTP 826
QY 113 TIPPSSPPQWKLOQDDPLLGNQ--SLLETHPLXOSEPAVPVTKTTP 157
Db 827 TSPVSTP---STAPAPSLAGDQQPSSAAFIPLISTRVSLRKTRQPP 870

RESULT 12

A39962
kinase-related transforming protein (abl) (EC 2.7.1.-) type I - mouse
N:Alternate names: protein-tyrosine kinase abl
C:Species: Mus musculus (house mouse)
C:Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 18-Jun-1999
C:Accession: A39962; A24773; S00771; A00626
R:Oppl, C.; Shore, S.K.; Reddy, E.P.
Proc. Natl. Acad. Sci. U.S.A. 84, 8200-8204, 1987
A:Title: Nucleotide sequence of testis-derived c-abl cDNAs: implications for testis-s
A:Reference number: A39962; MUID:88068561
A:Accession: A39962
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1123 <OPP>
A:Cross-references: GB:J02995; NID:g191566; PIDN:AAA88241.1; PID:g309084
A:Experimental source: clone AL6
R:Ben-Neriah, Y.; Bernards, A.; Paskind, M.; Daley, G.Q.; Baltimore, D.
Cell 44, 577-586, 1986
A:Title: Alternative 5' exons in c-abl mRNA.
A:Reference number: A90877; MUID:86133550
A:Accession: A24773
A:Molecule type: mRNA
A:Residues: 1-118 <BEN>
R:Bernards, A.; Paskind, M.; Baltimore, D.
Oncogene 2, 297-304, 1988
A:Title: Four murine c-abl mRNAs arise by usage of two transcriptional promoters and
A:Reference number: S00771; MUID:88202920
A:Accession: S00771
A:Molecule type: DNA
A:Residues: 1-26 <BER>
A:Cross-references: EMBL:X07539; NID:g49837; PIDN:CAA30411.1; PID:g49838
R:Wang, J.Y.J.; Ledley, F.; Goff, S.; Lee, R.; Groner, Y.; Baltimore, D.
Cell 36, 349-356, 1984
A:Title: The mouse c-abl locus: molecular cloning and characterization.
A:Reference number: A00626; MUID:84106840
A:Accession: A00626
A:Molecule type: DNA
A:Residues: 85-127, 'C', 129-182 <WAN>
C:Genetics:
A:Gene: abl
C:Superfamily: human protein-tyrosine kinase abl; protein kinase homology; SH2 homolo
C:Keywords: alternative splicing; ATP; autophosphorylation; nucleus; phosphoprotein;
F:68-116/Domain: SH3 homology <SH3>
F:127-217/Domain: SH2 homology <SH2>
F:240-500/Domain: protein kinase homology <KIN>
F:248-256/Region: protein kinase ATP-binding motif

```

Query Match          10.3%; Score 90; DB 2; Length 407;
Best Local Similarity 27.5%; Pred. No. 2.6;
Matches 44; Conservative 14; Mismatches 68; Indels 34; Gaps

Qy 11 VRRTE-GGEILLN---QLAGRMIPKGTLWS---CKFPTLDHVLHDVQTM---EEINT 58
Db 47 VARCESGNWSINTNGYLGLGLQTS--TWAHGGGEFAPSQLASREQQIANGERVLA 104
Qy 59 LQNOGAPGAGRVRGLSNTPTQEIPOPTPEDQKAREAFRRYOERPPETTTTPSS 118
Db 105 TQGRGAMPV-CGR--GLSNATPREVLPSAAMDAPLDAAV-----NGEPAPLAPPPAD 155
Qy 119 P-POWKLQPGDDPLLGNQSLLETHPLYQSEPAVPVTKTPP 157
Db 156 PAPPVELAANDLPA-----PLGEPLPAAPADAPP 185

RESULT 15
T16348
hypothetical protein F42G9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16348
R:Taich, A.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid F42G9.
A:Reference number: Z18498
A:Accession: T16348
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-283 <TAI>
A:Cross-references: EMBL:U00051; NID:g1216305; PID:g1216306; PIDN:AAA91952
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F42G9.3
A:Introns: 78/3; 179/1

Query Match          10.2%; Score 89; DB 2; Length 283;
Best Local Similarity 24.1%; Pred. No. 2;
Matches 40; Conservative 25; Mismatches 55; Indels 46; Gaps

Qy 12 RRIEGEILLNQLAGRMIPKGTLT--WSGKFEPTLDHVLHDVQWMEIN--TLNQGANPA 67
Db 51 RRVLSGE-----GRK-EKTKNLFLW-----ILTSRVDPETOYSQGSFPE 89
Qy 68 GAGRRVGLSNTPTQEIPOPTPEDQKAR-----EAFRVOERPPETTTIP- 115
Db 90 VAVPEPEPEPEPEPEPEPEPEPEPKPKKPEVDPVQPEPEYQPEPEPEPEVE 149
Qy 116 PSSPPQWKLQPGDDPLLGNQSLLETHPLYQSEPAVPVTKTPPLKK 161
Db 150 PEPEVPVPEPEPEP-----EPEVPVEKEEV-VVESPPREQE 186

Search completed: May 13, 2002, 11:22:32
Job time: 199 sec

```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:18:23 ; Search time 12.56 seconds
(without alignments)
288.458 Million cell updates/sec

Title: US-09-818-066-34

Perfect score: 871
Sequence: 1 MGQHPAKSMVRRIEGEIL.....PLYQSEPAVPVTKTPPLKKK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 871 | 100.0 | 161 | 2 | US-08-683-262B-34 |
| 2 | 871 | 100.0 | 161 | 4 | US-09-361-707-34 |
| 3 | 263 | 30.2 | 48 | 2 | US-08-683-262B-68 |
| 4 | 263 | 30.2 | 48 | 4 | US-09-361-707-68 |
| 5 | 93.5 | 10.7 | 906 | 2 | US-08-609-230A-9 |
| 6 | 93.5 | 10.7 | 907 | 3 | US-08-990-140-4 |
| 7 | 93.5 | 10.7 | 940 | 4 | US-08-810-712-7 |
| 8 | 88.5 | 10.2 | 1194 | 4 | US-08-538-526-1 |
| 9 | 88 | 10.1 | 16 | 2 | US-08-683-262B-18 |
| 10 | 88 | 10.1 | 16 | 4 | US-09-361-707-18 |
| 11 | 83.5 | 9.6 | 605 | 2 | US-08-687-956A-1 |
| 12 | 82.5 | 9.5 | 1257 | 1 | US-08-340-428B-49 |
| 13 | 82 | 9.4 | 214 | 1 | US-08-217-327-4 |
| 14 | 82 | 9.4 | 1664 | 2 | US-08-642-846-2 |
| 15 | 81.5 | 9.4 | 1199 | 1 | US-08-041-538-2 |
| 16 | 81.5 | 9.4 | 1199 | 1 | US-08-463-642-2 |
| 17 | 81.5 | 9.4 | 1199 | 1 | US-08-455-602-2 |
| 18 | 81.5 | 9.4 | 1199 | 2 | US-08-465-157-2 |
| 19 | 81.5 | 9.4 | 1199 | 5 | PCT-US91-09422-2 |
| 20 | 81.5 | 9.4 | 1219 | 2 | US-08-687-289A-6 |
| 21 | 79.5 | 9.1 | 667 | 4 | US-09-303-064-55 |
| 22 | 79.5 | 9.1 | 712 | 1 | US-08-587-889-2 |
| 23 | 79.5 | 9.1 | 712 | 2 | US-08-980-060-5 |
| 24 | 79.5 | 9.1 | 712 | 4 | US-09-307-185-5 |
| 25 | 79.5 | 9.1 | 712 | 5 | PCT-US96-09193-2 |
| 26 | 79 | 9.1 | 1004 | 4 | US-08-916-352-2 |
| 27 | 79 | 9.1 | 1128 | 1 | US-08-111-939-2 |

| | | | | | | |
|----|------|-----|------|---|-------------------|--------------------|
| 28 | 78.5 | 9.0 | 853 | 3 | US-09-254-325-2 | Sequence 2, Appli |
| 29 | 78.5 | 9.0 | 856 | 3 | US-09-079-415-6 | Sequence 6, Appli |
| 30 | 78.5 | 9.0 | 856 | 3 | US-08-750-458A-2 | Sequence 2, Appli |
| 31 | 78.5 | 9.0 | 1346 | 2 | US-08-635-121-2 | Sequence 2, Appli |
| 32 | 78 | 9.0 | 312 | 3 | US-08-894-017-10 | Sequence 10, Appli |
| 33 | 78 | 9.0 | 424 | 3 | US-09-173-581-7 | Sequence 7, Appli |
| 34 | 78 | 9.0 | 424 | 4 | US-09-420-915-7 | Sequence 23, Appli |
| 35 | 78 | 9.0 | 1561 | 3 | US-08-894-017-23 | Sequence 7, Appli |
| 36 | 77.5 | 8.9 | 180 | 4 | US-08-483-533-29 | Sequence 29, Appli |
| 37 | 77.5 | 8.9 | 355 | 4 | US-08-483-533-41 | Sequence 41, Appli |
| 38 | 77.5 | 8.9 | 355 | 5 | PCT-US91-06532-3 | Sequence 3, Appli |
| 39 | 77.5 | 8.9 | 396 | 3 | US-09-082-310-2 | Sequence 2, Appli |
| 40 | 77.5 | 8.9 | 550 | 3 | US-08-659-188-18 | Sequence 18, Appli |
| 41 | 77.5 | 8.9 | 550 | 3 | US-08-655-227-18 | Sequence 18, Appli |
| 42 | 77.5 | 8.9 | 550 | 3 | US-08-655-241-18 | Sequence 18, Appli |
| 43 | 77.5 | 8.9 | 746 | 3 | US-09-144-759-18 | Sequence 18, Appli |
| 44 | 77.5 | 8.9 | 764 | 3 | US-09-144-759-20 | Sequence 20, Appli |
| 45 | 77 | 8.8 | 276 | 3 | US-08-633-993A-15 | Sequence 15, Appli |

ALIGNMENTS

RESULT 1
US-08-683-262B-34
; Sequence 34, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-262B-34

Query Match 100.0%; Score 871; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGQHPAKSMVRRIEGEILLNOLAGRMIPKGLTWSGKPTLDHVLHDHVTMEINTLQ 60
Db 1 MGQHPAKSMVRRIEGEILLNOLAGRMIPKGLTWSGKPTLDHVLHDHVTMEINTLQ 60
QY 61 NOGAWPAGARRVGLSNPTPOEIPQWTPEDQKAREAFRRYQEEPRPETTTIPPSPP 120

```
Db 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPSSPP 120
|
QY 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
|
Db 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
|
RESULT 2
US-09-361-707-34
; Sequence 34, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-8906
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-361-707-34

Query Match 100.0%; Score 871; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQHPAKSMVDRRIEGEILLNLQAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|
Db 1 MGQHPAKSMVDRRIEGEILLNLQAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
|
QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPSSPP 120
|
Db 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPSSPP 120
|
QY 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
|
Db 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
|
RESULT 3
US-08-683-262B-68
```

```
; Sequence 68, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-683-262B-68

Query Match 30.2%; Score 263; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 AGRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPP 116
|
Db 1 AGRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPP 48
|
RESULT 4
US-09-361-707-68
; Sequence 68, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
```


APPLICATION NUMBER: 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Creason, Gary L.
REGISTRATION NUMBER: 34,310
REFERENCE/DOCKET NUMBER: 00786/287003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-361-707-68

Query Match 30.2%; Score 263; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 AGRVCLSNPTQEIPOQWTFEEDQKAREAFRRYQEEPPETTTIPP 116
Db 1 AGRVCLSNPTQEIPOQWTFEEDQKAREAFRRYQEEPPETTTIPP 48

RESULT 5
US-08-609-230A-9
; Sequence 9, Application US/08609230A
; Patent No. 5866333
; GENERAL INFORMATION:
; APPLICANT: Innerarity, Thomas L.
; APPLICANT: Qian, Xiaobing
; APPLICANT: Yamataka, Shinya
; TITLE OF INVENTION: Screening Methods to Detect mRNA Targets
; TITLE OF INVENTION: of Editing Enzymes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,230A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 023070-068100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 650-326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-230A-9

Query Match 10.7%; Score 93.5; DB 2; Length 906;
Best Local Similarity 25.3%; Pred. No. 0.14;

Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;
QY 1 MGQHPAKSMQVRRRIEGGTEIL-----LNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTM 54
Db 397 MGRH--RSNQLFNGHGGHIMPPTQSQFGEMGKFKMSQGLS-----OLYHNQSQG 444
QY 55 EINTLQNGAWPAGAGRRVGLSNPTQEIPOQWTFEE--DQKAREAFRRYQEEPPETTTIPP 110
Db 445 LLSQLQGGQS-----KDMPPRFSSKQGLNADEISLRPAQSFLMKNKQVPLQPOQ 492
QY 111 TTTIPPSS--PQWKLOP-GDDPLGNQSLLETH-PLYQSEPAVPVVIKTPPLKKK 161
Db 493 ITMIPPSAQPPRTQTPPLGQTPQLG----LKTNPPLIQEKPAKTSKKQPPPSKEE 542

RESULT 6
US-08-990-140-4
; Sequence 4, Application US/08990140A
; Patent No. 6093795
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Methot, Nathalie
; APPLICANT: Rom, Eran
; TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) and Human
; FILE REFERENCE: 1488.0700001
; CURRENT APPLICATION NUMBER: US/08/990,140A
; CURRENT FILING DATE: 1997-12-12
; EARLIER APPLICATION NUMBER: US 60/033,151
; EARLIER FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-990-140-4

Query Match 10.7%; Score 93.5; DB 3; Length 907;
Best Local Similarity 25.3%; Pred. No. 0.14;
Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;
QY 1 MGQHPAKSMQVRRRIEGGTEIL-----LNQLAGRMIPKGTLTWSGKFPPTLDHVDHVQTM 54
Db 398 MGRH--RSNQLFNGHGGHIMPPTQSQFGEMGKFKMSQGLS-----OLYHNQSQG 445
QY 55 EINTLQNGAWPAGAGRRVGLSNPTQEIPOQWTFEE--DQKAREAFRRYQEEPPETTTIPP 110
Db 446 LLSQLQGGQS-----KDMPPRFSSKQGLNADEISLRPAQSFLMKNKQVPLQPOQ 493
QY 111 TTTIPPSS--PQWKLOP-GDDPLGNQSLLETH-PLYQSEPAVPVVIKTPPLKKK 161
Db 494 ITMIPPSAQPPRTQTPPLGQTPQLG----LKTNPPLIQEKPAKTSKKQPPPSKEE 543

RESULT 7
US-08-810-712-7
; Sequence 7, Application US/08810712G
; Patent No. 6160106
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. LTD
; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
; TITLE OF INVENTION: Use of said genes and proteins
; FILE REFERENCE: sequence list
; CURRENT APPLICATION NUMBER: US/08/810,712G
; CURRENT FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/11598
; EARLIER FILING DATE: 1994-10-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1

QY 16 GGEILLNQLAGRMTPKGTLTWSGRFPPTLDHVID-HVQTMEETINTLQNCA-----WPA 67
| : :: :||| | :||| | :||| |
Matches 39; Conservative 23; Mismatches 52; Indels 49; Gaps 8;

```

RESULT 10
US-09-361-707-18
: Sequence 18, Application US/09361707
: Patent No. 6258937
: GENERAL INFORMATION:
: APPLICANT: Tong, Shuping
: Li, Jisu
: Wands, Jack R.
: TITLE OF INVENTION: HEPA DNA
: NUMBER OF SEQUENCES: 117
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson

```

RESULT 12
US-08-340-428B-49
; Sequence 49, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.428B
; FILING DATE: 14 No. 5648465ember 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/922,911
; FILING DATE: 03 August 1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: Margolis-1a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-428B-49

Query Match 9.5%; Score 82.5; DB 1; Length 1257;
Best Local Similarity 25.4%; Pred. No. 3.3;
Matches 33; Conservative 14; Mismatches 42; Indels 41; Gaps 5;

QY 37 SGKEPTLDHVDHV---QTMEENTLQNCWNP-----AGAGRVGLSNTPQPIQ 85
Db 506 SAQPTTLEVTADHGPSAATALESQSHSPWAILTNEVDVPGAGSLGSRSLPESR----- 561
QY 86 PWTPEEDQKAREAFRRYQERPPETTTIPSPSPQWK-----LQPGDDPLLGNSLLETH 141
Db 562 -KNSP-----SLSPSTVPSTDTSTFGLKPGADEAPGVKSAIHHP 599
QY 142 PLYOSEPAVP 151
Db 600 PWPSPPAVP 609

RESULT 13

US-08-217-327-4
; Sequence 4, Application US/08217327
; Patent No. 5474925
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal E
; APPLICANT: Barton, Kenneth A
; TITLE OF INVENTION: Immobilized Proteins in Cotton Fiber
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles and Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,327
; FILING DATE:
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/812,233
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 1122990831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-217-327-4

Query Match 9.4%; Score 82; DB 1; Length 214;
Best Local Similarity 31.2%; Pred. No. 0.33;
Matches 25; Conservative 2; Mismatches 41; Indels 12; Gaps 2;

QY 78 PTQEIPOQWTPBEDQKAREAFRRYQERPPETTTIPSPSPQWKLOPQDDPLLGNSL 137
Db 52 PTEPTATPPPVSTPTPTSSPPPV-----TASPPPVSTPTPTSSPPPVSTPTPTSSPPPV 99
QY 138 LETHPLYQSEPAVPVVIKTPP 157
Db 100 PATPPPPASPPPATPPPPASPP 119

RESULT 14

US-08-642-846-2
; Sequence 2, Application US/08642846
; Patent No. 5886151
; GENERAL INFORMATION:
; APPLICANT: HOSTETTER, MARGARET K.
; APPLICANT: GALE, CHERYL A.
; APPLICANT: BENDEL, CATHERINE M.
; APPLICANT: TAO, NIAN-JUN
; APPLICANT: KENDRICK, KATHLEEN
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,846
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00280101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-846-2

Query Match 9.4%; Score 82; DB 2; Length 1664;
Best Local Similarity 25.8%; Pred. No. 5.4;
Matches 42; Conservative 26; Mismatches 59; Indels 36; Gaps 10;

QY 3 QHPAKSMDVRRIEGGEILLNQLAGRMIPKGTLTWSGKFPTLDHVDL-HVQTMEEINTLQ 61
Db 99 QQFQQQQQLSQTNN--LIDEFSFQTPTMTSTLDTKONPTVDKVNENHAPTY--INTSPN 154
QY 62 QG-----AWPAGAGRRVG--LSNPTTPQEIPOQWTPBEDQKAREAFRRYQERPPETTTIP 115
Db 155 KSIKKATPKASPKKVAFTVTNPEIHHPDNR-VEEEDQSQK-----EDSVEP 202
QY 116 PSSPPQWKLOPQDDPLLGNSLLETHPLYQSEPAVPVVIKTPPL 158
Db 203 PLIQHONWK-----DPSQFNYSDEDTN-----ASVP--PTPPL 232

RESULT 15

US-08-041-538-2
 ; Sequence 2, Application US/08041538
 ; Patent No. 5385831
 ; GENERAL INFORMATION:
 ; APPLICANT: Mulvihill, Eileen R
 ; APPLICANT: Hagen, Frederick S
 ; APPLICANT: Houamed, Khaled M
 ; APPLICANT: Almers, Wolfhard
 ; TITLE OF INVENTION: G PROTEIN COUPLED GLUTAMATE
 ; TITLE OF INVENTION: RECEPTORS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend
 ; STREET: Steuart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/041,538
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/648,481
 ; FILING DATE:
 ; APPLICATION NUMBER: US 07/626,806
 ; FILING DATE: 12-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parmelee, Steven W
 ; REGISTRATION NUMBER: 31,990
 ; REFERENCE/DOCKET NUMBER: 13952-6-1
 ; TELEPHONE: 206-467-9600
 ; TELEFAX: 206-623-6793
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1199 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-041-538-2

Query Match 9.4%; Score 81.5; DB 1; Length 1199;
 Best Local Similarity 22.9%; Pred. No. 3.9;
 Matches 36; Conservative 19; Mismatches 65; Indels 37; Gaps 5;
 Qy 16 GGEILLNLAGRMIPKGTLTWSGRFPTLDHVDHVTMEETLONQCA-----WPAG 68
 Db 890 GKSWSWSEPGGRQAPKQGHVW-----QRLSVHVKTNE---TACNQTAIVIKPLTKSYQG 939
 Qy 69 AGRRVGLSNPTPQEIPOQWTPEDQKAREFRYOEERPPETTTIPSPSPQWKLOPG- 127
 Db 940 SGKSLTFS-----DASTKTLNVEEDNTPSAHFSPSPSMVVRGCP 983
 Qy 128 ---DDPLGNOSLLETHLYOSEPAVPVIKTPPLKKK 161
 Db 984 PVATTPLPPLHAEETPLFLADSVIPKGLPPLPQ 1020

Search completed: May 13, 2002, 11:20:18
 Job time: 115 sec

This Page Blank (uspto)

| Result No. | Query | | | Length | DB | ID | Description |
|------------|-------|-------|------|--------|----------|----|--------------------|
| | Score | Match | | | | | |
| 1 | 871 | 100.0 | 161 | 18 | AAW11854 | | Duck hepatitis B v |
| 2 | 95.5 | 11.0 | 424 | 21 | AAG31344 | | Arabidopsis thalia |
| 3 | 95.5 | 11.0 | 559 | 21 | AAG31343 | | Arabidopsis thalia |
| 4 | 95.5 | 11.0 | 633 | 21 | AAG31342 | | Arabidopsis thalia |
| 5 | 93.5 | 10.7 | 906 | 18 | AAW35681 | | Transgenic mouse N |
| 6 | 93.5 | 10.7 | 907 | 19 | AAW49032 | | Human eIF4G-like p |
| 7 | 93.5 | 10.7 | 940 | 19 | AAW71370 | | Death associated p |
| 8 | 93.5 | 10.7 | 940 | 21 | AAW58970 | | Breast and ovarian |
| 9 | 90 | 10.3 | 407 | 20 | AAW95400 | | M. tuberculosis RP |
| 10 | 90 | 10.3 | 407 | 22 | AAW52470 | | Mycobacterium tube |
| 11 | 89 | 10.2 | 1319 | 22 | AAW40120 | | Human polypeptide |

FT Peptide 25..102 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 42..102 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 71..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 71..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 80..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 80..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 87..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 87..104 /note= "claimed pre-S fragment (Claim 21)
FT capable of binding hepadnavirus receptor"
XX WO9704000-A1.
XX 06-FEB-1997.
XX 22-JUL-1996; 96WO-US12098.
XX 21-JUL-1995; 95US-0001371.
XX (GEHO) GEN HOSPITAL CORP.
XX Li J, Tong S, Wands JR;
XX WPI; 1997-132572/12.
XX N-PSDB; AAT59583.
XX p170, a new avian hepadnavirus receptor - binds to pre-S domain of
XX duck hepatitis B virus, attenuated p170 may be used to immunise
XX animals against hepadnaviral infection
XX Claim 19; Page 87-88; 175pp; English.
XX The pre-S domain (AAW11854) of the duck hepatitis B virus (DHBV)
XX large envelope protein binds to the cellular p170 receptor (see
XX also AAW11851-52) at a major neutralising epitope, within which are 2
XX basic amino acids (lys-95, Arg-97) required for virion-receptor
XX interaction. Peptide sequences encompassing lys-95 and Arg-97
XX can act as p170 ligands. Such peptides can be used to reduce
XX the level of hepadnaviral infection in an animal. The pre-S
XX domain (esp. attenuated), and nucleic acids encoding it, can also
XX be used as vaccines to immunise animals against hepadnavirus

CC infection.
XX Sequence 161 AA;
SQ
Query Match 100.0%; Score 871; DB 18; Length 161;
Best local Similarity 100.0%; Pred. No. 6e-76;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGQHPAKSMVRRIGEGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVTMEINTLQ 60
|||||
DB 1 mgqhpaksmvrrieggeillnqlagrmipkgtltwsghftldhvdhvtmeeintlq 60
|||||
QY 61 MGCAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQBERPPTTTPSSPP 120
|||||
DB 61 ngawpagarrvrglsnptqeiqpqwtpeedqkareafrryqgeerppettitppssp 120
|||||
QY 121 QWKLQPCDDPLGNQSLLETHPLYSQSEPAVPVTKTPPLKKK 161
|||||
DB 121 qwklqpgddplgnqsllethplysqsepavpvtktpplkkk 161
RESULT 2
AAG31344
ID AAG31344 standard; Protein; 424 AA.
XX AAG31344;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37626.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 20-MAY-1999; 99US-0146386.
PR 21-MAY-1999; 99US-0146388.
PR 24-MAY-1999; 99US-0146389.
PR 25-MAY-1999; 99US-0147038.
PR 27-MAY-1999; 99US-0147204.
PR 28-MAY-1999; 99US-0147302.
PR 01-JUN-1999; 99US-0147302.
PR 03-JUN-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 07-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.0%; Score 95.5; DB 21; Length 424;
Best Local Similarity 24.5%; Pred. No. 0.48;
Matches 40; Conservative 15; Mismatches 45; Indels 63; Gaps 10;

QY 21 LNLGGRMIPKGLT--WSGKFTPLDHLVDH-----VQTWEINTLQNCQAWPA 67
Db 264 vqellgk---tgnvtsfw-----mdhikthcyvyspsveeaaatreavynlq----wpp 310
QY 68 GAGRRV-----GLSNFTPOEIQPOWTPEDOKAREAFRRYQEEPPETTIP 115
Db 311 nggrhlaefvraeevkeleaplp--pqpqhqpqaqlslr-----ppptalp 356
QY 116 PSSPPQWKLPQGGDPLLGNSLLETHPLVQSEPAVPVVIKTPPL 158
Db 357 p--pp-----plakpphvverlplpppppiabeeqeppl 388

RESULT 3
AAG31343
ID AAG31343 standard; Protein; 559 AA.
XX
AC AAG31343;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37625.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
PE 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144352.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.
99US-0145087.
99US-0145089.
99US-0145192.
99US-0145145.
99US-0145218.
99US-0145224.
99US-0145276.
99US-0145913.
99US-0145918.
99US-0145919.
```


PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.0%; Score 95.5; DB 21; Length 633;
Best Local Similarity 24.5%; Pred. No. 0.79;
Matches 40; Conservative 15; Mismatches 45; Indels 63; Gaps 10;

Qy 21 LNQLAGRMIPKGTLT--WSGKFPFLDHLVDH-----VOTMEINTLQOGAWPA 67
Db 473 vqellgk---tgnvtstf-----mdhikthcyvspsveaaatreaavnlq-----wpp 519

Qy 68 GAGRRV-----GLSNPTQEIPOQWTPPEEDOKAREAFRRYQERPPETTTIP 115
Db 520 nggrhllaeafvraeevkeleaplp---pqhqhpqatlsr-----ppptaip 565

Qy 116 PSSPPQWKLOPDDPLLGNSLLETHPLYQSEPAVPVVIKTPPL 158
Db 566 p--pp-----plakpphvvverlpplppppiapeegeppi 597

RESULT 5
AAW25681
ID AAW25681 standard; Protein; 906 AA.
AC AAW25681;
XX
DT 31-MAR-1998 (first entry)
XX
DE Transgenic mouse NTA1 polypeptide.
XX
KW RNA editing; mooring primer; screening; detection; editing enzyme;
KW obesity; APOBEC-1; therapeutic; NTA1; transgenic.
XX
OS Synthetic.
OS Homo sapiens.
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Protein 1..906
FT /label= NTA1
FT /note= "partial coding sequence"
XX
PN WO9732032-A1.
XX
PD 04-SEP-1997.
XX
PF 28-FEB-1997; 97WO-US03189.
XX
PR 01-MAR-1996; 96US-0609230.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Innerarity TL, Qian X, Yamanaka S;
XX
WPI: 1997-448694/41.
DR N-PSDB; AAT86087.
DR
XX
PT Detection of mRNA targets for editing enzymes - by amplifying RNA
PT from tissues by RT-PCR using mooring primers to identify genes
PT responsible for non-wild type phenotype(s)
XX
PS Example 4; Figure 5; 36pp; English.
XX
CC This sequence represents a novel target of the apoB mRNA-editing
CC enzyme catalytic polypeptide #1, APOBEC-1, (i.e NTA1) which is used
CC as an example of a novel method of detecting mRNA candidates for editing
CC in a tissue. The method involves the generation of cDNA from RNA in the
CC tissue using one or more mooring primers and reverse transcriptase (RT)
```

```
CC and is amplified by PCR using one or more mooring primers and a 5'
CC arbitrary primer. The method has been used for identification of an
CC obesity gene e.g. APOBEC-1 or an oncogene in an animal and also for the
CC identification of non-wild type phenotypes. The identification of such
CC mRNA's provides insight into the role of the genes encoding them, in
CC growth, differentiation, or lipid accumulation, and also into the
CC possible role of an RNA editing enzyme or polypeptide in causing obesity
CC or as an oncogene. The identified polypeptides can be used as therapeutic
CC reagents in a pathological state. Additionally, such polypeptides or the
CC edited versions can be used in the preparation of antibodies for
CC therapeutic use.
XX
SQ Sequence 906 AA;

Query Match 10.7%; Score 93.5; DB 18; Length 906;
Best Local Similarity 25.3%; Pred. No. 1.9;
Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;

Qy 1 MGQHPAKSMDEVRRIEGGEIL-----LNQLAGRMIPKGTLTWSGKFPFLDHLVDHVTME 54
Db 397 mgrh--rsnqlfngghimpptqsgfgemgkfkmsqgls-----qlyhnqsgg 444

Qy 55 EINTLQOGAWPAGRRVGLSNPTQEIPOQWTPPEE-DOKAREAFRRYQERP---PE 110
Db 445 llsqlqggs-----kdmpprfskkgqlnadeislrpaqsfImnknqvpklqpg 492

Qy 111 TTTTPSS-PPQWKLOP-GDDPLLGNSLLETH-PLYQSEPAVPVVIKTPPLKKK 161
Db 493 itmppsagqprrtqpplgqtpqig----lknppllqekpaktskkpppskee 542

RESULT 6
AAW49032
ID AAW49032 standard; Protein; 907 AA.
XX
AC AAW49032;
XX
DT 07-OCT-1998 (first entry)
XX
DE Human eIF4G-like protein (p97).
XX
KW Human Prt1-like subunit protein; hPrt1; transcription; apoptosis;
KW p97; antagonist; agonist; human eIF4G-like protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 788..802
FT /note= "INT peptide used to raise anti-INT
FT antibodies"
XX
PN WO9825957-A2.
XX
PD 18-JUN-1998.
XX
PF 12-DEC-1997; 97WO-US22664.
XX
PR 13-DEC-1996; 96US-0033151.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (UYMC-) UNIV MCGILL.
XX
PI Inataka H, Methot N, Olsen HS, Rom E, Ruben SM;
PI Sonenberg N;
XX
WPI: 1998-348451/30.
DR N-PSDB; AAV32791.
DR
XX
PT Polynucleotides encoding human hprt1 and p97 - useful for treating
PT disease states associated with apoptosis
XX
```



```
XX Gennaro ML, Gomez MJ;
XX WPI; 2001-007151/01.
XX
XX Novel Mycobacterium tuberculosis secreted polypeptides and
XX polynucleotides useful in diagnosis, treatment and prophylaxis of
XX tuberculosis -
XX
XX Claim 11; Fig 1; 60pp; English.
XX
XX The present invention relates to Mycobacterium tuberculosis secreted
XX proteins (MTPSP), where the polypeptide has M. tuberculosis specific
XX antigenic and immunogenic properties. Compositions of the invention may
XX be useful for diagnosing Mycobacterium tuberculosis infection and as a
XX vaccine against M. tuberculosis infection.
XX
XX Sequence 407 AA;
SQ
Query Match 10.3%; Score 90; DB 22; Length 407;
Best Local Similarity 27.5%; Pred. No. 1.5;
Matches 44; Conservative 14; Mismatches 68; Indels 34; Gaps 10;
QY 11 VRRIE-GGEILLN---QLAGRMIPKCTLTWS-----GKFTLDHVLHVQTM---ESINT 58
Db 47 varcesggnwsintngylgglgtqgs--twaahgggefapsaqilasredqgiavgervla 104
QY 59 LQNOGAMPAGAGRRVGLSNPTPOIPQWPTPEEDQKAREAFRRYQERPPPTTTIPSS 118
Db 105 tqgrgawpv-cgr--glisnatprevlpasaamdapladaav-----ngepaplappad 155
QY 119 P-PQWKLPQGDPLGLNQSLLETHPLYSQSEPAVPVKTTP 157
Db 156 pappvelaandlpa-----plgeplpaapadpapp 185
RESULT 11
AAM40120
ID AAM40120 standard; Protein; 1319 AA.
XX
XX AAM40120;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 3265.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0862191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX N-PSDB; AA159276.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 5; SEQ ID NO 3265; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 1319 AA;
SQ
Query Match 10.2%; Score 89; DB 22; Length 1319;
Best Local Similarity 28.6%; Pred. No. 8.4;
Matches 32; Conservative 9; Mismatches 41; Indels 30; Gaps 4;
QY 64 ANPAGAGRRVGLSN---PTQEIPOPOWPTPEEDQKAREAFRRYQ----- 104
Db 276 ampraag-mvglskmhagppqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq 334
QY 105 -----EERPPTTTTPSSPPQWKLPQGDPLGLNQSLLETHPLYSQSEPA 149
Db 335 hplmqppqpppppppppppppppppppppppppppppppppppppppppppppppppp 383
RESULT 12
AAR42199
ID AAR42199 standard; Protein; 1194 AA.
XX
XX AAR42199;
XX
XX 16-MAY-1994 (first entry)
XX
XX HSmGluR1.
XX
XX Human; metabotropic; glutamate receptor; HSmGluR1; drug; stroke;
XX Alzheimer's disease; head trauma; epilepsy.
XX
XX Homo sapiens.
XX
XX EP569240-A.
XX
XX 10-NOV-1993.
XX
XX 06-MAY-1993; 93EP-0303520.
XX
XX 08-MAY-1992; 92US-0884571.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Burnett JP, Mayne NG, Sharp RL, Snyder YM;
XX WPI; 1993-353316/45.
XX
```


DR N-PSDB: AAQ50879.
XX New human metabotropic glutamate receptor protein - and related
PT DNA, vectors and expression systems, used to screen drugs for
PT treatment of neurological disease
XX
XX Claim 1: Page 16-19; 31pp; English.
PS
XX This sequence represents the human metabotropic glutamate receptor
CC protein (Hsmglu1). This protein may be used for screening drugs
CC which modulate the receptor, eg. drugs which may be useful for
CC treating stroke, Alzheimer's disease, head trauma and epilepsy. To
CC identify compounds which interact with, or effect the receptor,
CC cells expressing it are contacted with the test substance and any
CC interaction measured.
XX
XX Sequence 1194 AA;
SQ
Query Match 10.2%; Score 88.5; DB 14; Length 1194;
Best Local Similarity 23.9%; Pred. No. 8.2;
Matches 39; Conservative 23; Mismatches 52; Indels 49; Gaps 8;
Qy 16 GGEILLNQLAGRMIPKGLTWSGKFEPTLDHVID-HVQTMREINTLQNGA-----WPA 67
Db 890 gksvswsepggvgpdkghw-----hrslsvhvtne---tacnqtavikpltksyq 938
Qy 68 GAGRRVGLSNPTPOEIPQWTPEDOKAREAFRRYQ-----EERPPETTTIPPSS 118
Db 939 gsgskltsdstktl-----ynveeedaqpl--rfspgsgpsmvhrrvpsaattpplp 992
Qy 119 PPQWKLPQDDPLLGNSLLETHPLYOSEPAVPVIKTPPLKKK 161
Db 993 p-----hitaeetplflaepalpglpplqqg 1020
RESULT 13
AAM38840
ID AAM38840 standard; Protein; 396 AA.
AC AAM38840;
XX
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 1985.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB: AAI57996.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Example 3; SEQ ID NO 1985; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 396 AA;
SQ
Query Match 10.0%; Score 87.5; DB 22; Length 396;
Best Local Similarity 30.6%; Pred. No. 2.6;
Matches 30; Conservative 10; Mismatches 33; Indels 25; Gaps 3;
Qy 67 AGAGR-----RVGLSNPTPOEIPQW--TPEDOKAREAFRRYQEE 106
Db 18 agagrfciglseaatrkhlparnhcglsdsspglwpdpdfrnpprkaskasaldfkryvtd 77
Qy 107 RPPTTTI-----PPSPPPQWKLPQDDPLLGNSLLE 139
Db 78 rrlaetlaqlylgkpsrpphlllecnpgpgiltqalle 115
RESULT 14
AAM40626
ID AAM40626 standard; Protein; 407 AA.
XX AAM40626;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 5557.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.

```
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0862191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AAI59782.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Example 2; SEQ ID NO 5557; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAI38642-AAI42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 407 AA;
SQ

Query Match 10.0%; Score 87.5; DB 22; Length 407;
Best Local Similarity 30.6%; Pred. No. 2.7;
Matches 30; Conservative 10; Mismatches 33; Indels 25; Gaps 3;

QY 67 AGAGR-----RVGLSNTPQEIPOQW--TPEDQKAREAFRRYQEE 106
Db 29 agagfcilgseatrhlparnhcglsdsspglwpdpfrnpprkaskasldfkrytd 88
QY 107 RPPETTTI-----PPSSPPQWKLPQGDPLLGNSLLE 139
Db 89 rrlaetlaqlylgkpsrpphlliecnpgpgiltqalle 126

RESULT 15
AAB23501
ID AAB23501 standard; Protein; 528 AA.
XX
XX AAB23501;
XX
XX 05-JAN-2001 (first entry)
XX
XX Human negative elongation factor (NELF) protein SEQ ID 2.
DE
XX
XX Negative elongation factor; NELF; RNA polymerase II; pol II; human;
KW cytostatic; cancer; transcription elongation inhibition.
XX
XX Homo sapiens.
XX
XX JP2000189168-A.
PN
XX
XX 11-JUL-2000.
PD
XX
XX 28-DEC-1998; 98JP-0373999.
XX
XX
```

```
PR 28-DEC-1998; 98JP-0373999.
XX (HAND/) HANDA H.
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX WPI: 2000-545765/50.
DR N-PSDB; AAA90082.
XX
XX A new transcription elongation inhibiting factor -
PT
XX Claim 4; Page 22-24; 45pp; Japanese.
XX
XX This sequence represents a human negative elongation factor (NELF)
CC protein. NELF inhibits the elongation reaction of RNA polymerase II, pol
CC II, in the presence of 5,6-dichloro-beta-D-ribofuranosylbenzimidazole
CC (DRB) a sensitivity inducing factor. Sequences AAA90082-A90085 and
CC AAB23501-B23505 represent NELF DNA and protein sequences, while
CC AAB23506-B23519 represent NELF peptides. The invention includes
CC recombinant NELF DNA, a transformant carrying the recombinant DNA,
CC antibodies recognising the NELF protein, and a method for detecting the
CC protein using the antibodies. NELF exhibits cytostatic activity, and the
CC proteins, polynucleotides and peptides can be used in the treatment of
CC cancers.
XX
XX Sequence 528 AA;
SQ

Query Match 9.9%; Score 86.5; DB 21; Length 528;
Best Local Similarity 27.1%; Pred. No. 4.6;
Matches 35; Conservative 17; Mismatches 52; Indels 25; Gaps 8;

QY 42 TLDHVLHDHVTMEINTLQNGAWPAGARRVGLSNTPQEIPOQWTPEDQKAREAFR 101
Db 297 tpdyaaglvst-qklgslnnpalps-----tsylptpsvvpssyipsstetppapssr 350
QY 102 RYQEEERPETTTP-PSSPPQWK-----LQPG-DDPLLGNQSLLETHPLYQSEPAV 150
Db 351 --easrppeepsapsptipagfkqrapmynsgispatptpaaptspitptp-----pav 403
QY 151 -PVIKTPPL 158
Db 404 apttqtpv 412

Search completed: May 13, 2002, 11:19:59
Job time: 211 sec
```